EXHIBIT

Digital Video and HDTV Algorithms and Interfaces

Charles Poynton

MORGAN KAUFMANN PUBLISHERS
AN IMPRINT OF ELSEVIER SCIENCE
AMSTERDAM BOSTON LONDON NEW YORK
OXFORD PARIS SAN DIEGO SAN FRANCISCO
SINGAPORE SYDNEY TOKYO

Publishing Director: Diane Cerra
Publishing Services Manager: Edward Wade
Production Editor: Howard Severson
Design, illustration, and composition: Charles Poynton
Editorial Coordinator: Mona Buehler
Cover Design: Frances Baca
Copyeditor: Robert Fiske
Proofreader: Sarah Burgundy
Printer: The Maple-Vail Book Manufacturing Group
Cover images: Close-up of woman/Eyewire; Circuit connection/Artville; Medical prespectives/Picture Quest

Morgan Kaufmann Publishers
An imprint of Elsevier Science
340 Pine Street, Sixth Floor
San Francisco, CA 94104-3205
www.mkp.com

Printed in the United States of America

2007 2006 2005 2004 2003 5 4 3 2 1

Library of Congress Control Number: 2002115312
ISBN: 1-55860-792-7

This book is printed on acid-free paper.

Compression ratio	Quality/application	Example tape formats
2:1	"Visually lossless" studio video	Digital Betacam
3.3:1	Excellent-quality studio video	DVCPRO50, D-9 (Digital-S)
6.6:1	Good-quality studio video; consumer digital video	D-7 (DVCPRO), DVCAM, consumer DVC, Digital8

Table 14.2 **Approximate compression ratios of M-JPEG** for SDTV applications

MPEG

Apart from scene changes, there is a statistical likelihood that successive pictures in a video sequence are very similar. In fact, it is *necessary* that successive pictures are similar: If this were not the case, human vision could make no sense of the sequence!

The M in MPEG stands for *moving*, not *motion*!

M-JPEG's compression ratio can be increased by a factor of 5 or 10 by exploiting the inherent temporal redundancy of video. The *MPEG* standard was developed by the *Moving Picture Experts Group* within ISO and IEC. In MPEG, an initial, self-contained picture provides a base value – it forms an *anchor* picture. Succeeding pictures can then be coded in terms of pixel differences from the anchor, as sketched in Figure 14.1 at the top of the facing page. The method is termed *interframe coding* (though differences between *fields* may be used).

Once the anchor picture has been received by the decoder, it provides an estimate for a succeeding picture. This estimate is improved when the encoder transmits the prediction errors. The scheme is effective provided that the prediction errors can be coded more compactly than the raw picture information.

Motion may cause displacement of scene elements – a fast-moving element may easily move 10 pixels in one frame time. In the presence of motion, a pixel at a certain location may take quite different values in successive pictures. Motion would cause the prediction error information to grow in size to the point where the advantage of interframe coding would be negated.

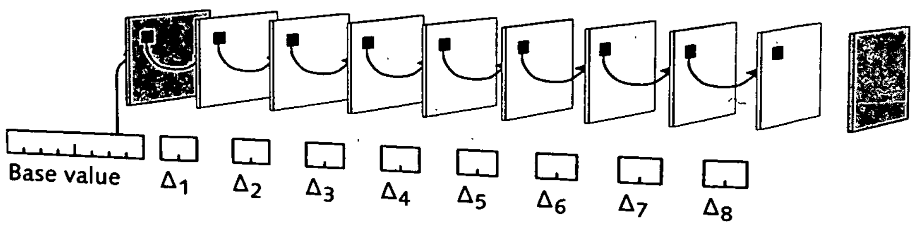


Figure 14.1 **Interpicture coding** exploits the similarity between successive pictures in video. First, a base picture is transmitted (ordinarily using intra-picture compression). Then, pixel differences to successive pictures are computed by the encoder and transmitted. The decoder reconstructs successive pictures by accumulating the differences. The scheme is effective provided that the difference information can be coded more compactly than the raw picture information.

However, objects tend to retain their characteristics even when moving. MPEG overcomes the problem of motion between pictures by equipping the encoder with *motion estimation* circuitry: The encoder computes *motion vectors*. The encoder then displaces the pixel values of the anchor picture by the estimated motion – a process called *motion compensation* – then computes prediction errors from the motion-compensated anchor picture. The encoder compresses the prediction error information using a JPEG-like technique, then transmits that data accompanied by motion vectors.

Based upon the received motion vectors, the decoder mimics the motion compensation of the encoder to obtain a predictor much more effective than the undisplaced anchor picture. The transmitted prediction errors are then applied to reconstruct the picture.

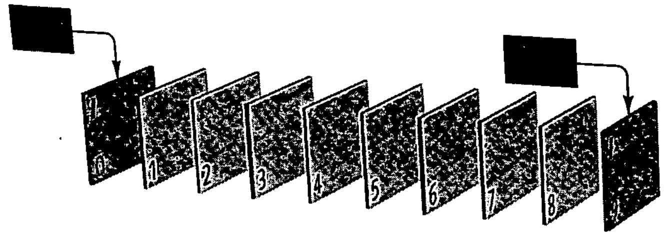
When encoding interlaced source material, an MPEG-2 encoder can choose to code each field as a picture or each frame as a picture, as I will describe on page 478. In this chapter, and in Chapter 40, the term *picture* can refer to either a field or a frame.

Picture coding types (I, P, B)

In MPEG, a video sequence is typically partitioned into successive *groups of pictures* (GOPs). The first frame in each GOP is coded independently of other frames using a JPEG-like algorithm; this is an *intra picture* or *I-picture*. Once reconstructed, an I-picture becomes an anchor picture available for use in predicting neighboring (*nonintra*) pictures. The example GOP sketched in Figure 14.2 overleaf comprises nine pictures.

A *P-picture* contains elements that are predicted from the most recent anchor frame. Once a P-picture is reconstructed, it is displayed; in addition, it becomes

Figure 14.2 **MPEG group of pictures (GOP).** The GOP depicted here has nine pictures, numbered 0 through 8. I-picture 0 is decoded from the coded data depicted in the dark gray block. Picture 9 is not in the GOP; it is the first picture of the next GOP. Here, the *intra count* (*n*) is 9.



a new anchor. I-pictures and P-pictures form a two-layer hierarchy. An I-picture and two dependent P-pictures are depicted in Figure 14.3 below.

MPEG provides an optional third hierarchical level whereby *B-pictures* may be interposed between anchor pictures. Elements of a B-picture may be bidirectionally predicted by averaging motion-compensated elements from the past anchor and motion-compensated elements from the future anchor. Each B-picture is reconstructed, displayed, and discarded: No B-picture forms the basis for any prediction. (At the encoder's discretion, elements of a B-picture may be unidirectionally forward-interpolated from the preceding anchor, or unidirectionally backward-predicted from the following anchor.) Using B-pictures delivers a substantial gain in compression efficiency compared to encoding with just I- and P-pictures.

Two B-pictures are depicted in Figure 14.4 at the top of the facing page. The three-level MPEG picture hierarchy is summarized in Figure 14.5 at the bottom of the facing page; this example has the structure IBBPBBPBB.

Figure 14.3 **An MPEG P-picture** contains elements forward-predicted from a preceding anchor picture, which may be an I-picture or a P-picture. Here, the first P-picture (3) is predicted from an I-picture (0). Once decoded, that P-picture becomes the predictor for the second P-picture (6).

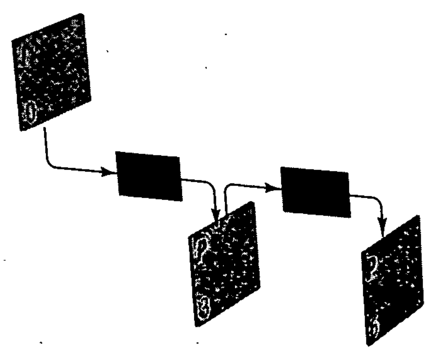
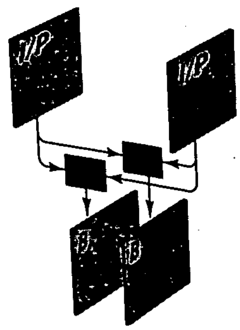
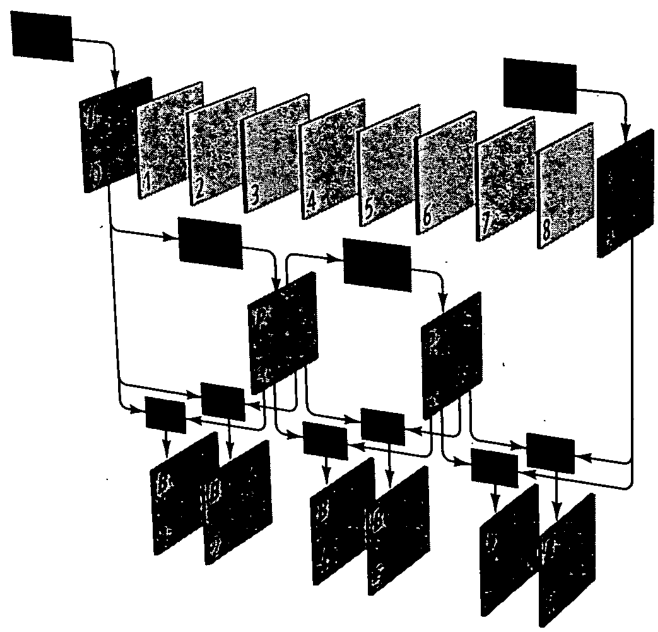


Figure 14.4 **An MPEG B-picture** is generally estimated from the average of the preceding anchor picture and the following anchor picture. (At the encoder's option, a B-picture may be unidirectionally forward-predicted from the preceding anchor, or unidirectionally backward-predicted from the following anchor.)



A simple encoder typically produces a bitstream having a fixed schedule of I-, P-, and B-pictures. A typical GOP structure is denoted IBBPBBPBBPBBPBB. At 30 pictures per second, there are two such GOPs per second. Regular GOP structure is described by a pair of integers *n* and *m*; *n* is the number of pictures from one I-picture (inclusive) to the next (exclusive), and *m* is the number of pictures from one anchor picture (inclusive) to the next (exclusive). If $m = 1$, there are no B-pictures. Figure 14.5 shows a regular GOP structure with an I-picture interval of $n = 9$ and an anchor-picture interval of $m = 3$. The $m = 3$ component indicates two B-pictures between anchor pictures.

Figure 14.5 **The three-level MPEG picture hierarchy.** This sketch shows a regular GOP structure with an I-picture interval of n=9, and an anchor-picture interval of m=3. This example represents a simple encoder that emits a fixed schedule of I-, B-, and P-pictures; this structure can be described as IBBPBBPBB. This example depicts an *open GOP*, where B-pictures following the last P-picture of the GOP are permitted to use backward prediction from the I-frame of the following GOP. Such prediction precludes editing of the bitstream between GOPs. A *closed GOP* permits no such prediction, so the bitstream can be edited between GOPs.



Coded B-pictures in a GOP depend upon P- and I-pictures; coded P-pictures depend upon earlier P-pictures and I-pictures. Owing to these interdependencies, an MPEG sequence cannot be edited, except at GOP boundaries, unless the sequence is decoded, edited, and subsequently reencoded. MPEG is very suitable for distribution, but owing to its inability to be edited without impairment at arbitrary points, MPEG is unsuitable for production. In the specialization of MPEG-2 called *I-frame only MPEG-2*, every GOP is a single I-frame. This is conceptually equivalent to Motion-JPEG, but has the great benefit of an international standard. (Another variant of MPEG-2, the *simple profile*, has no B-pictures.)

I have introduced MPEG as if all elements of a P-picture and all elements of a B-picture are coded similarly. But a picture that is generally very well predicted by the past anchor picture may have a few regions that cannot effectively be predicted. In MPEG, the image is tiled into *macroblocks* of 16×16 luma samples, and the encoder is given the option to code any particular macroblock in *intra* mode – that is, independently of any prediction. A compact code signals that a macroblock should be *skipped*, in which case samples from the anchor picture are used without modification. Also, in a B-picture, the encoder can decide on a macroblock-by-macroblock basis to code using forward prediction, backward prediction, or bidirectional prediction.

Reordering

In a sequence without B-pictures, I- and P-pictures are encoded and transmitted in the obvious order. However, when B-pictures are used, the decoder typically needs to access the past anchor picture and the future anchor picture to reconstruct a B-picture.

Figure 14.6 **Example GOP**

$\mathbf{I_0}B_1B_2\mathbf{P_3}B_4B_5\mathbf{P_6}B_7B_8$

Consider an encoder about to compress the sequence in Figure 14.6 (where anchor pictures I_0, P_3, and P_6 are written in boldface). The coded B_1 and B_2 pictures may be backward predicted from P_3, so the encoder must buffer the uncompressed B_1 and B_2 pictures until P_3 is coded: Only when coding of P_3 is complete can coding of B_1 start. Using B-pictures incurs a penalty in

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☑ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☑ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: ____________________

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.